

SKRIPSI

ANALISIS DAN KLASIFIKASI SENTIMEN TERHADAP

TWITTER STMIK AKAKOM YOGYAKARTA

MENGGUNAKAN NAÏVE BAYES CLASSIFIER



YUNA SOPHIA DEWI FEBRIANT

Nomor Mahasiswa : 135410076

PROGRAM STUDI TEKNIK INFORMATIKA

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN

KOMPUTER AKAKOM YOGYAKARTA

2017

SKRIPSI

ANALISIS DAN KLASIFIKASI SENTIMEN TERHADAP
TWITTER STMIK AKAKOM YOGYAKARTA
MENGGUNAKAN NAIVE BAYES CLASSIFIER

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana
Komputer Program Studi Teknik Informatika
Sekolah Tinggi Manajemen Informatika dan Komputer
Akakom Yogyakarta

Disusun Oleh :

YUNA SOPHIA DEWI FEBRIANT

Nomor Mahasiswa : 135410076

PROGRAM STUDI TEKNIK INFORMATIKA
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN
KOMPUTER AKAKOM
YOGYAKARTA

2017

LEMBAR PERSETUJUAN

Judul : Analisis dan Klasifikasi Sentimen Terhadap Twitter
STMIK AKAKOM Yogyakarta Menggunakan Naïve Bayes
Classifier

Nama : Yuna Sophia Dewi Febriant

NIM : 135410076

Jurusan : Teknik Informatika (S1)

Semester : Genap 2016/2017

Telah memenuhi syarat dan disetujui untuk di selenggarakan dihadapan dosen
penguji seminar tugas akhir

Yogyakarta, 24 Agustus 2017

Dosen Pembimbing



Sri Redjeki, S.Si., M.Kom

HALAMAN PENGESAHAN

ANALISIS DAN KLASIFIKASI SENTIMEN TERHADAP TWITTER STMIK AKAKOM YOGYAKARTA MENGGUNAKAN NAIVE BAYES CLASSIFIER

Telah diuji didepan Dosen Penguji Skripsi dan dinyatakan diterima sebagai syarat
untuk memperoleh gelar Sarjana Komputer di Sekolah Tinggi Manajemen

Informatika dan

Komputer Akakom Yogyakarta

Yogyakarta, 24 Agustus 2017

Mengesahkan,

Dosen Penguji:

Deborah Kurniawati, S.Kom., M.Cs.

Dini Fakta Sari, S.T., M.T.

Sri Redjeki, S.Si., M.Kom.

Mengetahui,

Ketua Jurusan Teknik Informatika

24 AUG 2017



Ir. M. Guntara, M.T.

HALAMAN PERSEMBAHAN



Puji syukur kepada Allah SWT atas segala rahmat dan karunianya. Kerja keras, niat, ketekunan dan kebanggaan yang saya curahkan menjadi karya tulis ini, saya persembahkan kepada :

Keluarga, yang cinta dan do'anya tiada habis untuk saya, Abah, Mama dan Okta. Tidaklah Allah ridho kepada saya tanpa ada ridho dari kalian. Terimakasih atas rasa cinta dan do'anya yang menjelma menjadi keajaiban, kebaikan, semangat dan segala sesuatu yang tumbuh bersama tumbuhnya putri kalian ini.

Ibu Sri Redjeki, S.Si., M.Kom., dan Ibu Dini Fakta Sari, S.T., M.T., selaku dosen pembimbing yang telah memberikan bimbingan, pengarahan, semangat, waktu, dan ilmunya kepada saya dalam penyusunan Tugas Akhir ini

Teman-teman yang telah banyak memberikan masukan dan semangat dalam pengerjaan skripsi ini yang tidak bisa saya sebutkan satu persatu namanya didalam naskah tugas akhir saya. Dan yang selalu memberikan semangat dan waktunya, Alif Yudhia Pratama.

MOTTO

“Bekerjalah bagaikan tak butuh uang. Mencintailah bagaikan tak pernah disakiti.

Menarilah bagaikan tak seorang pun sedang menonton”

-Mark Twain-

INTISARI

Makin maraknya penggunaan jejaring sosial seperti Twitter menjadikan jejaring sosial tersebut sebagai data yang sangat besar. Salah satu pemanfaatan data ini adalah untuk mengetahui opini atau sentimen pengguna jejaring sosial terhadap suatu topik. Keberadaan Twitter telah digunakan secara luas oleh berbagai lapisan masyarakat dalam beberapa tahun terakhir. Oleh karena itu, penelitian ini mencoba mencoba menganalisis *tweet* berbahasa Indonesia yang membicarakan tentang STMIK Akakom Yogyakarta.

Analisis dilakukan dengan melakukan klasifikasi *tweet* yang berisi opini masyarakat tentang STMIK Akakom Yogyakarta. Klasifikasi sentimen terdiri dari positif, netral dan negatif. Penelitian ini menerapkan proses *text mining* serta menggunakan algoritma Naïve Bayes Classifier untuk mengklasifikasi sentimen dari *tweet* tersebut. Penelitian ini menggunakan 1000 data *tweet* yang diperoleh dari Twitter API untuk keperluan data *training*. Data tersebut diklasifikasi secara manual untuk menentukan sentimen dari *tweet* tersebut. Kemudian 50 data *tweet* digunakan untuk *testing*.

Hasil penelitian ini menghasilkan sebuah sistem yang dapat mengklasifikasi sentimen secara otomatis dengan hasil pengujian 78% untuk sentimen, beberapa kendala dalam pemrosesan tweet secara realtime seperti duplikasi tweet (tweet spam), struktur Bahasa Indonesia yang cukup rumit dan beragam, dan belum banyak penelitian atau algoritma untuk optimasi pada kalimat berbahasa Indonesia.

Kata Kunci : *Analisis Sentimen, Analisis Sentimen dan Klasifikasi, Naive Bayes Classifier, Twitter.*

KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa, berkat kasih dan karuni-Nya sehingga penulis dapat menyelesaikan Tugas Akhir ini yang berjudul Analisis dan Klasifikasi Sentimen Terhadap Twitter STMIK AKAKOM Yogyakarta Menggunakan Naive Bayes Classifier.

Dalam penulisan Tugas Akhir ini tentunya tidak lepas dari bantuan berbagai pihak, oleh karena itu pada kesempatan ini penulis ingin menyampaikan rasa terimakasih kepada semua pihak yang telah membantu dalam penulisan tugas akhir ini, antara lain :

1. Bapak Cuk Subiantoro, S.Kom., M.Kom., selaku Ketua Sekolah Tinggi Manajemen Informatika dan Komputer AKAKOM Yogyakarta.
2. Bapak Ir. Guntara, M.T. selaku Ketua Jurusan Teknik Informatika Jengjang Strata Satu (S1) di Sekolah Tinggi Manajemen Informatika dan Komputer AKAKOM Yogyakarta yang telah memberikan saya kesempatan dan kepercayaan untuk menyelesaikan Tugas Akhir tepat waktu.
3. Ibu Sri Redjeki, S.Si., M.Kom., selaku dosen pembimbing yang telah memberikan bimbingan, pengarahan, semangat, waktu, dan ilmunya kepada saya dalam penyusunan Tugas Akhir ini.
4. Ibu Dini Fakta Sari, S.T., M.T., dan Ibu Deborah Kurniawati, S.Kom., M.Cs., selaku dosen narasumber yang telah banyak memberikan bimbingan dan arahan kepada saya dalam pengerjaan metode pada tugas akhir yang saya buat.

5. Seluruh dosen dan staff karyawan Sekolah Tinggi Manajemen Informatika dan Komputer AKAKOM Yogyakarta.
6. Terimakasih kepada Orang Tua, Keluarga, dan Teman-teman saya, yang telah banyak memberikan masukan dan semangat dalam pengerjaan tugas akhir ini.

Penulis menyadari bahwa Tugas Akhir ini jauh dari kata sempurna, oleh karena itu semua kritik dan saran yang bersifat membangun sangatlah diharapkan. Semoga hasil karya tulis ini bisa memberikan manfaat bagi semua pihak.

Yogyakarta,

2017

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERSEMBAHAN	iv
MOTTO	v
INTISARI.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABLE	xiii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Ruang Lingkup.....	3
1.4. Tujuan Penelitian	4
1.5. Manfaat Penelitian	5
1.6. Sistematika Penelitian	5
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI	6
2.1. Tinjauan Pustaka	6
2.2. Dasar Teori	9
2.2.1. Twitter	9
2.2.2. Twitter API.....	9
2.2.3. Text Mining	11
2.2.4. Sentiment Analysis	12
2.2.5. Naïve Bayes Classifier.....	13
BAB III METODE PENELITIAN	17
3.1. Analisis Kebutuhan.....	17
3.1.1. Kebutuhan Data	17
3.1.2. Kebutuhan Proses	17

3.1.3.	Kebutuhan Perangkat Lunak	18
3.1.4.	Kebutuhan Perangkat Keras	18
3.2.	Perancangan Sistem	19
3.2.1.	Blok Diagram	19
3.2.2.	Usecase Diagram	22
3.2.3.	Sequence Diagram	23
3.2.4.	Class Diagram	24
3.2.5.	Activity Diagram	25
3.2.6.	Perancangan Antarmuka	26
BAB IV	IMPLEMENTASI DAN PEMBAHASAN	32
4.1.	Implementasi Sistem	32
4.1.1.	Algoritma Analisis dan Klasifikasi	32
4.1.2.	Crawling Data	33
4.1.3.	Text Preprocessing	33
4.1.4.	Training Data	44
4.1.5.	Testing Data	49
4.2.	Pembahasan Sistem	50
4.2.1.	Pembahasan Halaman Database	50
4.2.2.	Pembahasan Halaman Login	56
4.2.3.	Pembahasan Halaman Admin	57
4.2.4.	Meminta dan Menampilkan Data pada Aplikasi Client	61
4.2.5.	Pembahasan Sistem	64
BAB V	KESIMPULAN	65
5.1.	Kesimpulan	65
5.2.	Saran	65
DAFTAR PUSTAKA	67
LAMPIRAN	68

DAFTAR GAMBAR

Gambar 3.1. Blok Diagram Sistem.....	19
Gambar 3.2. UseCase Diagram Sistem	22
Gambar 3.3. Sequence Diagram Sistem	23
Gambar 3.4. Class Diagram Sistem.....	24
Gambar 3.5. Activity Diagram Sistem	25
Gambar 3.6. Rancangan Halaman Login.....	27
Gambar 3.7. Rancangan Halaman Administrator.....	27
Gambar 3.8. Rancangan Halaman Headline.....	28
Gambar 3.9. Rancangan Halaman Stopword.....	29
Gambar 3.10. Rancangan Halaman Dictionary	29
Gambar 3.11. Rancangan Halaman Utama Client.....	30
Gambar 3.12. Rancangan Halaman Pengecekan	31
Gambar 3.13. Rancangan Grafik Analisis	31
Gambar 4.1. Potongan program get data twitter.....	33
Gambar 4.2. Potongan Script Case Folding dan Remove Punctuation	34
Gambar 4.3. Potongan script Tokenizing	35
Gambar 4.4. Potongan script Stopword Removal	35
Gambar 4.5. Potongan script Stemming	36
Gambar 4.6. Halaman Database Controller.....	51
Gambar 4.7. Halaman Database Table Admin	51
Gambar 4.8. Halaman Database Table Dataheadline	52
Gambar 4.9. Halaman Database Table Dictionary	54
Gambar 5.10. Halaman Database Table Stopword.....	55
Gambar 4.11. Halaman Database Table word_count	56
Gambar 4.12. Halaman Login Admin	57
Gambar 4.13. Halaman Home Admin	58
Gambar 4.14. Halaman Headline	58
Gambar 4.15. Halaman Form Data Headline	59

Gambar 4.16. Halaman Stopword	60
Gambar 4.17. Halaman Dictionary	61
Gambar 4.18. Halaman Utama Client	61
Gambar 4.19. Halaman get twit.....	62
Gambar 4.20. Halaman Pengecekan Manual.....	63
Gambar 4.21. Halaman Grafik Analisis	63

DAFTAR TABLE

Tabel 2.1. Tinjauan Pustaka	8
Tabel 2.2. Tinjauan Pustaka (Lanjutan).....	9
Tabel 4.1. Gambaran contoh kasus <i>training</i>	45
Tabel 4.2. Perhitungan Probabilitas Training(1)	46
Tabel 4.3. Perhitungan Probabilitas Training(2)	46
Tabel 4.4. Penentuan Kategori Headline	46
Tabel 4.5. Nilai probabilitas data Testing	49